

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:)
)
 Eiichi TAKAHASHI, et al.)
) Group Art Unit: TBA
 Serial No.: TBA)
) Examiner: TBA
 Filed: September 28, 2000)



For: **SERVICE DISTRIBUTION DEVICE**

**SUBMISSION OF CERTIFIED COPY OF PRIOR FOREIGN
APPLICATION IN ACCORDANCE
WITH THE REQUIREMENTS OF 37 C.F.R. §1.55**

*Assistant Commissioner for Patents
Washington, D.C. 20231*

Sir:

In accordance with the provisions of 37 C.F.R. §1.55, the applicants submit herewith a certified copy of the following foreign application:

Japanese Patent Application No. 279516
Filed: September 30, 1999

It is respectfully requested that the applicant(s) be given the benefit of the foreign filing date as evidenced by the certified papers attached hereto, in accordance with the requirements of 35 U.S.C. §119.

Respectfully submitted,
STAAS & HALSEY LLP

By: C. Joan Gilsdorf
Christine Joan Gilsdorf
Registration No. 43,635

700 11th Street, N.W., Ste. 500
Washington, D.C. 20001
(202) 434-1500
Date: 9/28/00

日 本 国 特 許 庁
PATENT OFFICE
JAPANESE GOVERNMENT

Jc932 U.S. PTO
09/671120
09/28/00

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日

Date of Application:

1999年 9月30日

出 願 番 号

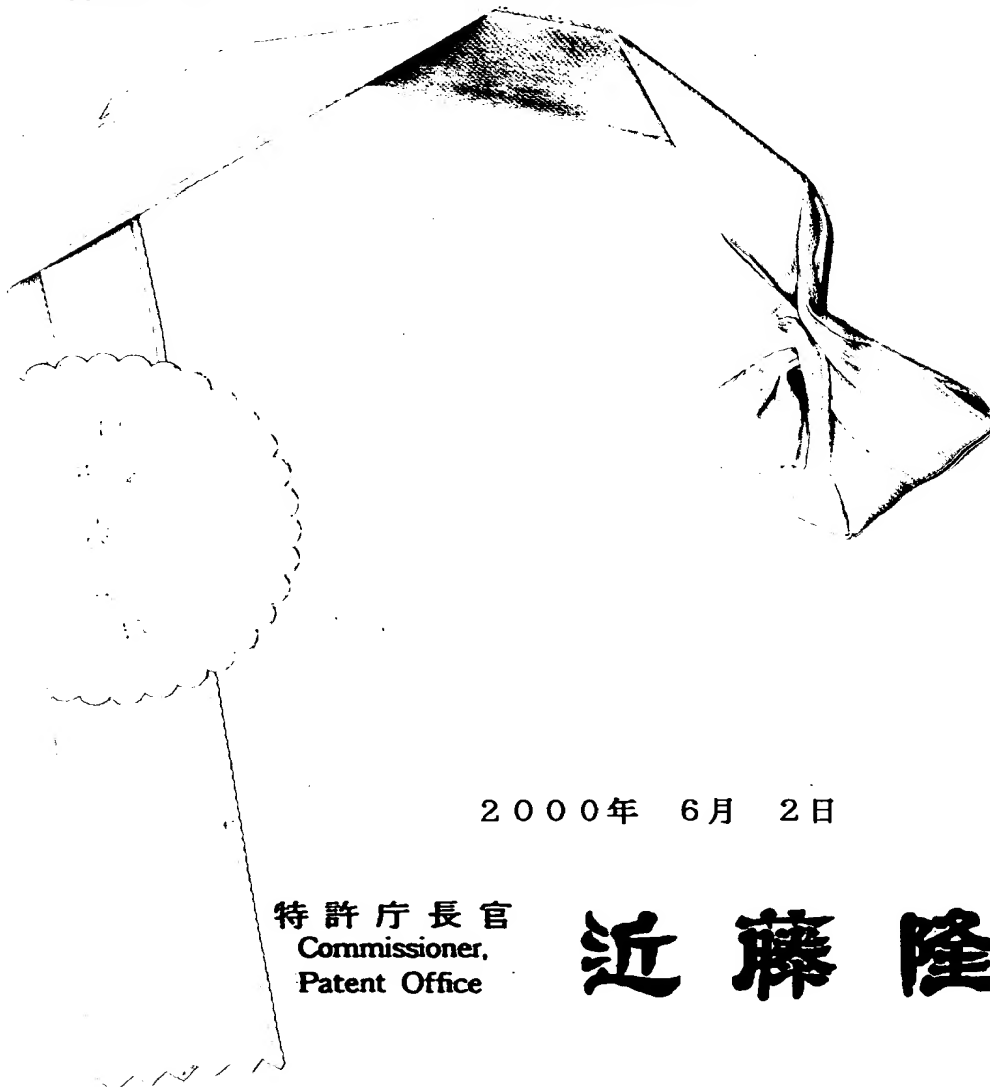
Application Number:

平成11年特許願第279516号

出 願 人

Applicant (s):

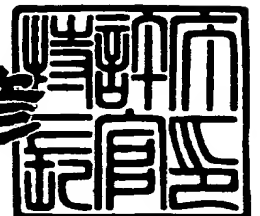
富士通株式会社



2000年 6月 2日

特許庁長官
Commissioner,
Patent Office

近 藤 隆 彦



出証番号 出証特2000-3040457

【書類名】 特許願

【整理番号】 9902570

【提出日】 平成11年 9月30日

【あて先】 特許庁長官殿

【国際特許分類】 H04L 29/00

【発明の名称】 サービス振り分け装置

【請求項の数】 11

【発明者】

 【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

 【氏名】 高橋 英一

【発明者】

 【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

 【氏名】 横山 乾

【発明者】

 【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

 【氏名】 田村 直広

【特許出願人】

 【識別番号】 000005223

 【氏名又は名称】 富士通株式会社

【代理人】

 【識別番号】 100072590

 【弁理士】

 【氏名又は名称】 井桁 貞一

 【電話番号】 044-754-3035

【手数料の表示】

 【予納台帳番号】 011280

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9704486

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 サービス振り分け装置

【特許請求の範囲】

【請求項 1】 サービスを複数台のサーバへサーバ負荷がバランスするように振り分けるために、サービスの振り分け先を決定するサービス振り分け装置であって、

ネットワークを流れるパケットをキャプチャするパケットキャプチャと、

キャプチャしたパケットに関するパケット情報をサーバ別にサーバログ部に記録するサーバ弁別部と、

キャプチャしたパケットに関するパケット情報をサービス別にサービスログ部に記録するサービス弁別部と、

サーバ別の記録からサーバについてのシミュレーション・モデルを設定するサーバモデリング部と、

サービス別の記録からサービスについてのシミュレーション・モデルを設定するサービスモデリング部と、

サーバモデルとサービスモデルを読み込み、シミュレーションを行うシミュレータと、

シミュレータの結果から最適な振り分け先サーバを選択し、指定するサーバ選択部とから構成されることを特徴とするサービス振り分け装置。

【請求項 2】 クライアントとサーバ間のパケットを中継するパケット中継装置上にパケットキャプチャが設けられ、パケット中継装置からパケットを取得することを特徴とする請求項 1 に記載のサービス振り分け装置。

【請求項 3】 前記サーバモデリング部は、送信処理に対応する待ち行列を持つモデルとしてサーバモデルを構成し、サーバ送信スループット、サーバ処理時間および単位処理時間をパラメータとし、

サーバログの記録に基づき、

連続送信されたパケット列の任意の連続部分列についての、(a) 先頭パケットキャプチャ時刻 $t-s$ 、(b) 末尾パケットキャプチャ時刻 $t-e$ 、および(c) その部分列で先頭パケットを除いたサイズの合計 L から、

式 $L / (t_e - t_s)$ を用いて、サーバ送信スループットを求め、

(d) サーバ応答パケットのキャプチャ時刻 t_s 、サイズ l_s 、(e) 対応するクライアント応答パケットのキャプチャ時刻 t_c 、サイズ l_c 、および (f) ネットワーク速度 B から、

式 $(t_s - t_c) - (l_s + l_c) / B$ を用いて、サーバ処理時間を求めることを特徴とする請求項 1 に記載のサービス振り分け装置。

【請求項 4】 前記サービスモデリング部は、

(a) 全サービスのセッション数に占める各サービスのセッション数の割合、
(b) セッションの開始頻度あるいは時間間隔、(c) セッション当たりのクライアントーサーバ間の送受信回数、(d) 送受信当たりのクライアント応答のサイズ、パケットサイズ、パケット数、(e) 送受信当たりのサーバ応答のサイズ、パケットサイズ、パケット数、および (f) サーバ応答からクライアント応答までの時間をパラメータとして、サービス毎にサービスモデルを構成し、

サービスログの記録から、それらのパラメータを求めることを特徴とする請求項 1 に記載のサービス振り分け装置。

【請求項 5】

前記シミュレータは、サーバモデルとサービスモデルを用いてシミュレーションを行ない、指定サービスについてセッション時間の平均値あるいは中間値を結果とすることを特徴とする請求項 1 に記載のサービス振り分け装置。

【請求項 6】 前記サーバ選択部において、サービス別に前記シミュレータで 1 セッションのみをシミュレートした結果を該サービスの基準値とし、複数セッションについてのシミュレーション結果値と該基準値の比率あるいは差が予め設定した閾値を超えれば高負荷と判断することを特徴とする請求項 1 に記載のサービス振り分け装置。

【請求項 7】 前記サーバ選択部において、振り分け先サーバ問い合わせを受けたとき、サーバ毎に対象サービスについて高負荷となるセッション開始頻度を所定の探索手法により求め、それを該サーバの許容度とし、現時点の対象サービスのセッション開始頻度と許容度との差が最も大きなサーバを振り分け先サーバとして指定することを特徴とする請求項 6 に記載のサービス振り分け装置。

【請求項 8】 前記サーバ選択部において、振り分け先サーバ問い合わせを受けたとき、サーバ毎に対象サービスについてシミュレーションを行ない、結果が基準値の β （所定値）倍となる割合が γ （所定値）以下となるサーバを振り分けサーバとして指定することを特徴とする請求項 6 に記載のサービス振り分け装置。

【請求項 9】 前記サーバ選択部において、振り分け先サーバ問い合わせを受けたとき、サーバ毎に対象サービスについてシミュレーションを行ない、結果が基準値の β （所定値）倍となる割合が最も小さいサーバを振り分け先サーバとして指定することを特徴とする請求項 6 に記載のサービス振り分け装置。

【請求項 10】 前記サービスモデリング部は、セッションの各送受信それぞれの処理内容に応じて、(a) 接続要求と応答、(b) コマンド送信、データ送信と応答、終了などに分類し、それぞれについてパラメータを求めることを特徴とする請求項 4 に記載のサービス振り分け装置。

【請求項 11】 前記サービス選択部は、サーバ毎に対象サービスについて高負荷となるセッション開始頻度を所定の探索手法により求め、それをサーバの許容度とし、(a) 各サーバの許容度をサービス振り分けの重み付けの値とするか、あるいは、(b) 許容度の相対比率をサーバ分配比率とすることを特徴とする請求項 7 に記載のサービス振り分け装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、クライアントからのネットワークを介したサービス要求を複数台のサーバに適切に振り分けるサービス振り分け装置に関する。

【0002】

【従来技術】

近年、ネットワークサービスの複雑化、大規模化に伴う高いサーバ性能への要求に対し、複数台のサーバへサービスを分散させ、台数分の性能を実現する負荷分散技術が一般に用いられている。台数分の性能を実現するためには、各サーバへそれぞれの処理性能に応じた量のサービスを振り分けて負荷をバランスする技

術が求められている。

【0003】

従来、いかに示す技術が知られている。

【0004】

- ・サーバのレスポンス時間をとらえ、最もレスポンス時間が短いサーバへサービスを振り分ける方式
- ・サーバが過負荷時（処理落ちが発生する状態）もしくは高負荷時（待ちが生じることでサービス時間が長くなる状態）のレスポンス時間と現在のレスポンス時間から最も過負荷もしくは高負荷になりにくいと予測したサーバへ振り分ける方式

【0005】

【発明が解決しようとする課題】

しかしながら、従来方式では次のような問題があった。

【0006】

- ・サーバ間に処理能力に差がある場合、最も高速なサーバに負荷が偏る、あるいはサービスの処理内容に差がある場合、サーバ負荷予測を誤り負荷バランスに失敗する。

【0007】

- ・各サーバがいずれも少なくとも一度は過負荷もしくは高負荷状態になれば閾値がわからず振り分けができない。閾値をユーザが予め設定することは困難である。意図的に閾値を求めるには、サービスを停止してベンチマークテストなどを実行しなければならない。

【0008】

【課題を解決するための手段】

本発明は、前記課題を解決するために、以下の構成を採用するものである。

【0009】

すなわち、請求項1の発明によるサービス振り分け装置は、ネットワークを流れるパケットをキャプチャするパケットキャプチャと、キャプチャしたパケットに関するパケット情報をサーバ別にサーバログ部に記録するサーバ弁別部と、キ

ャプチャしたパケットに関するパケット情報をサービス別にサービスログ部に記録するサービス弁別部と、サーバ別の記録からサーバについてのシミュレーション・モデルを設定するサーバモデリング部と、サービス別の記録からサービスについてのシミュレーション・モデルを設定するサービスモデリング部と、サーバモデルとサービスモデルを読み込み、シミュレーションを行うシミュレータと、シミュレータの結果から最適な振り分け先サーバを選択し、指定するサーバ選択部とから構成される。このような構成により、サーバやサービスの制限を受けず、サービストラフィックに影響を与えず、サーバを高負荷あるいは過負荷にすることなく自動的に各サーバの負荷をバランスさせる振り分け先サーバを決定することができる。

【0010】

請求項2の発明は、クライアントとサーバ間のパケットを中継するパケット中継装置上にパケットキャプチャが設けられ、パケット中継装置からパケットを取得する構成を採る。この構成により、パケットを取りこぼすことがなく確実にキャプチャでき、サーバとサービスモデルの精度を上げることができる。

【0011】

請求項3の発明は、請求項1のサーバモデリング部において、送信処理に対応する待ち行列を持つモデルとしてサーバモデルを構成し、サーバ送信スループット、サーバ処理時間および単位処理時間をパラメータとし、

サーバログの記録に基づき、

連続送信されたパケット列の任意の連続部分列についての、(a) 先頭パケットキャプチャ時刻 $t-s$ 、(b) 末尾パケットキャプチャ時刻 $t-e$ 、および(c) その部分列で先頭パケットを除いたサイズの合計 L から、

式 $L / (t-e - t-s)$ を用いて、サーバ送信スループットを求め、

(d) サーバ応答パケットのキャプチャ時刻 t_s 、サイズ l_s 、(e) 対応するクライアント応答パケットのキャプチャ時刻 t_c 、サイズ l_c 、および(f) ネットワーク速度 B から、

式 $(t_s - t_c) - (l_s + l_c) / B$ を用いて、サーバ処理時間を求める構成を採る。この構成により、サーバの仕様の詳細を知ることなく、サーバの正確なモデ

ルを設定することができる。

【 0 0 1 2 】

請求項 4 の発明は、請求項 1 のサービスモデリング部において、

- (a) 全サービスのセッション数に占める各サービスのセッション数の割合、
- (b) セッションの開始頻度あるいは時間間隔、(c) セッション当たりのクライアントーサーバ間の送受信回数、(d) 送受信当たりのクライアント応答のサイズ、パケットサイズ、パケット数、(e) 送受信当たりのサーバ応答のサイズ、パケットサイズ、パケット数、および(f) サーバ応答からクライアント応答までの時間をパラメータとして、サービス毎にサービスモデルを構成し、

サービスログの記録から、それらのパラメータを求める構成を採る。

この構成により、サービスの仕様の詳細や実装の相違を知ることなく、サービスの正確なモデルを設定できる。

【 0 0 1 3 】

請求項 5 の発明は、請求項 1 のシミュレータにおいて、サーバモデルとサービスモデルを用いてシミュレーションを行ない、指定サービスについてセッション時間の平均値あるいは中間値を結果とする構成を採る。

この構成により、クライアントから見たサービス品質を忠実に評価することができる。

【 0 0 1 4 】

請求項 6 の発明は、請求項 1 のサーバ選択部において、サービス別に前記シミュレータで 1 セッションのみをシミュレートした結果を該サービスの基準値とし、複数セッションについてのシミュレーション結果値と該基準値の比率あるいは差が予め設定した閾値を超えれば高負荷と判断する構成を採る。

この構成により、シミュレーション結果について自動的に負荷の高低を判断することができる。

【 0 0 1 5 】

請求項 7 の発明は、請求項 6 のサーバ選択部において、振り分け先サーバ間い合わせを受けたとき、サーバ毎に対象サービスについて高負荷となるセッション開始頻度を所定の探索手法により求め、それを該サーバの許容度とし、現時点の

対象サービスのセッション開始頻度と許容度との差が最も大きなサーバを振り分け先サーバとして指定する構成を採る。

この構成により、最もリソースに余裕があるサーバを少ない手間で選択することができる。

【0016】

請求項8の発明は、請求項6のサーバ選択部において、振り分け先サーバ問い合わせを受けたとき、サーバ毎に対象サービスについてシミュレーションを行ない、結果が基準値の β （所定値）倍となる割合が γ （所定値）以下となるサーバを振り分けサーバとして指定する構成を採る。

この構成により、振り分けてもサービス品質を落とさない可能性が最も高いサーバを選択することができる。

【0017】

請求項9の発明は、請求項6のサーバ選択部において、振り分け先サーバ問い合わせを受けたとき、サーバ毎に対象サービスについてシミュレーションを行ない、結果が基準値の β （所定値）倍となる割合が最も小さいサーバを振り分け先サーバとして指定する構成を採る。

この構成により、振り分けてもサービス品質を落とさない可能性が高いサーバの中で最も資源に余裕があるサーバを選択することができる。

【0018】

請求項10の発明は、請求項4のサービスモデリング部において、セッションの各送受信それぞれの処理内容に応じて、(a) 接続要求と応答、(b) コマンド送信、データ送信と応答、終了などに分類し、それぞれについてパラメータを求める構成を採る。

この構成により、サービスモデルの精度を上げることができる。

【0019】

請求項11の発明は、請求項7のサーバ選択部において、サーバ毎に対象サービスについて高負荷となるセッション開始頻度を所定の探索手法により求め、それをサーバの許容度とし、(a) 各サーバの許容度をサービス振り分けの重み付けの値とするか、あるいは、(b) 許容度の相対比率をサーバ分配比率とする構成を

採る。

この構成により、重み付けラウンドロビンまたはサービス分配比率に従い負荷を分散する装置における重み付けの値またはサービス分配比率を、実際の振り分けによる試行なしで求めることができ、試行ミスによるシステムのダメージやサービス品質低下を回避することができる。

【 0 0 2 0 】

【発明の実施の形態】

(実施例 1)

図 1 は、本発明によるサーバ振り分け装置 1 のネットワークシステム内の位置を示す図である。サーバ振り分け装置 1 は、ネットワーク 2 へサービスを行うサーバ # 1 ~ # Ns と同一セグメントに存在する。サーバ振り分け装置 1 は、負荷分散装置 3 から、パケットをどのサーバに振り分けるべきかという、振り分け先サーバの問い合わせを受けると、サーバ # 1 ~ # Ns の中から負荷をバランスさせるために適切なサーバを選択し、負荷分散装置 3 へ指定する。負荷分散装置 3 は、各種のネットワークとサーバ間のパケットを中継するものである。

【 0 0 2 1 】

図 1 に示す本発明のサーバ振り分け装置 1 の構成を図 2 に示す。サーバ振り分け装置は、ネットワーク 1 0 に接続しており、ネットワーク 1 0 を流れるパケットをパケットキャプチャ 1 1 でモニタする。パケットキャプチャ 1 1 はモニタしたパケットをサーバ弁別部 1 2 とサービス弁別部 1 3 へ送る。

【 0 0 2 2 】

サービス弁別部 1 3 は、パケットがどういうサービスのパケットであるかを調べる。例えば、IP パケットのヘッダから、サーバアドレスが送信アドレスにあれば送信番号を、宛て先アドレスにあれば受信ポート番号を読み取り、ポート番号をサービスの識別子としてよい。

【 0 0 2 3 】

サービス弁別部 1 3 は、サービスログ部 1 4 内にサービス毎に用意したログへ該パケット情報を記録する。パケット情報は例えば、時刻、パケットサイズ、クライアントアドレス、クライアントポート番号、シーケンス番号、SYN や AC

Kなどのフラグ、データの一部であってよい。

【0024】

サーバ弁別部12は、例えばIPパケットのヘッダから送信アドレスもしくは宛て先アドレスを読み取り、該パケットがサーバ#1～#Nsのいずれから発信されたものか、あるいはいずれへ発信されたものかを調べる。サーバ弁別部12はサーバを特定したら該パケットについての情報、例えばモニタした時刻、パケットサイズなどを、サーバログ部15内に用意したログへ記録する。

【0025】

パケットキャプチャ11、サービス弁別部13、サーバ弁別部12は常時動作してもよく、外部からの指示で動作/停止を行ってもよい。

[サービスモデルのパラメータ設定]

サービスモデリング部16は、サービスログ部14の記録からサービスについてのシミュレーション・モデルのパラメータをサービス毎に設定する。

パラメータは、例えば、

- ・サーバ毎に、ある1つのサービスの全サービス数に占める割合、
(例えば、A、B、C、Dの4つのサービスが存在する場合、Aの割合、Bの割合、Cの割合、Dの割合を求める)
- ・リクエスト(セッション)の頻度あるいは時間間隔、
- ・セッション当たりのクライアントとサーバ間での送受信回数、
- ・1回の送受信中のクライアントからの送信データサイズ、パケットサイズ、パケット数、
- ・1回の送受信中のサーバからの送信データサイズ、パケットサイズ、パケット数、
- ・サーバからのパケットに対するクライアントの応答までの時間、

などである。

【0026】

リクエスト頻度Fは例えば、時刻TxからTy(>Tx)の間にモニタしたサービス開始パケット数Pnから

$$F = P_n / (T_y - T_x)$$

で求める。

【0027】

クライアントとサーバ間の送受信回数 N_c は例えば、サービス開始から終了までのパケット数から求める。

【0028】

サーバからのパケットに対するクライアントの応答までの時間およびクライアントからのパケットに対するサーバの応答までの時間は、図4(a)に示す方法で求める。

【0029】

図4(a)はクライアントとサーバ間のサービス1つについての1セッション（開始から終了まで）のパケット送受信の例である。図中、モニタはパケットキャプチャ11（図2）によるパケットのモニタを意味し、パケット送受信を示す横向き（正確には、斜め下向き）の矢印と時間経過を示す下向きの矢印との交点は、パケットをキャプチャした時刻を意味する。

図4において、 t_i はパケットタイムスタンプ、 t_d_i は転送時間、 t_t_i は送信時間、 t_p_i はサーバ処理時間、 t_cr_i はクライアント応答時間である。

図4(a)の例では、サービスログ部14に、6つのパケット（開始、応答1～3、終了）についての、モニタした時刻 $t_1 \sim t_5$ と各パケットのサイズなどが記録される。

【0030】

クライアントの応答時間 t_cr_1 は、次式で求める。

【0031】

$$\begin{aligned} t_cr_1 &= (t_j - t_i) + t_d_j + t_d_i \\ t_d_i &= l_i / B \\ t_d_j &= l_j / B \end{aligned}$$

ここで、 t_i 、 t_j は対応するパケットのモニタ時刻、 l_i 、 l_j はパケットサイズ、 B はネットワーク速度であり、これらは実測により求められるも

のである。 $t d_i$ 、 $t d_j$ 、 $t c r_1$ は計算により求められるものである。

【0032】

上記パラメータそれぞれについて、値をログから複数求めてそれらの平均値としてよく、中間値としてもよく、確率分布を求めて確率変数としてもよい。

【0033】

サービスモデリング部 16 は、設定したモデルをサービスモデル部 17 へ出力する。

[サーバモデルのパラメータ設定]

次に、サーバモデリング部 18 は、サーバログ部 15 の記録からサーバ毎にサーバモデルのパラメータを設定する。パラメータは例えば、

- ・サーバ送信スループット α 、
- ・サーバ処理時間 $t p_i$ 、
- ・単位処理時間

などである。

【0034】

サーバ処理時間 $t p_i$ と送信時間 $t t_i$ を求める例を図 4 (a) に示す。

【0035】

サーバ送信スループット α はネットワーク速度 B と同一としてもよく、設定した値でもよい。図 4 (b) に示すように連続してサーバがパケットを送信するケースでは次式で α を求めてもよい。

【0036】

$$\alpha = L / (t 7 - t 6)$$

$$L = \sum l_j i - l_j 1$$

ここで、 $t 6$ は連続パケットの先頭パケットモニタ時刻、 $t 7$ は末尾パケットモニタ時刻、 L は連続パケットの全サイズから先頭パケットサイズを引いた値である。連続したパケットの連続した一部を用いてもよい。

【0037】

サーバ処理時間 t_{p_i} は次式で求める。

【0038】

$$t_{p_i} = (t_{_j} - t_{_i}) - t_{d_i} - t_{t_j}$$

$$t_{d_i} = l_{_i} / B$$

$$t_{t_j} = l_{_j} / B$$

ここで、 $t_{_i}$ はクライアントからのパケット i のモニタ時刻、 $t_{_j}$ はそのパケット i に対応するサーバからのパケット j のモニタ時刻である。

t_{d_i} はモニタからサーバまでのパケット i の転送時間であり、クライアントからのパケット i のサイズ $l_{_i}$ をネットワーク速度 B で割って得られるものである。

t_{t_j} はサーバからモニタまでのパケット j の送信時間であり、サーバからのパケット j のサイズ $l_{_j}$ をネットワーク速度 B で割って得られるものである。

【0039】

パラメータは平均値でよく、確率分布でもよく、設定値でもよい。

【0040】

サーバモデリング部 18 は、設定したモデルをサーバモデル部 19 へ出力する。

【0041】

サービス、サーバ両パラメータについて、サービスログの記録全てから求めてもよく、前回のモデリングで用いた記録以降から求めてもよい。また、時間間隔を一定となるようにしてもよい。Web サービスなどのように、サービス中の送受信それぞれを処理内容に応じ、

- ・ 接続要求と接続確立応答、
- ・ コマンド送信、データ送信と応答、終了

などに特徴付けることができるサービスについては、処理内容別に上記パラメータを求めてもよい。送受信の単位を例えば、TCP ならばデータ送信と ACK とし、シーケンス番号でそれらの対を識別してもよい。例えば、図 4 (b) で応答 j_1 、 j_2 と応答 k を送受信 1 回分としてよい。

サービスモデリング部 16、サーバモデリング部 18 は常時動作してもよく、外部からの指示で動作／停止してもよく、定期的に動作してもよい。

【0042】

シミュレータ 20 は、サービスモデル部 17 とサーバモデル部 19 からモデルを取り出し、処理 1（図 5 に示す）、処理 2（図 6 に示す）、処理 3（図 7）、セッションレベル処理（図 8 に示す）、サービスレベル処理 a（図 12 に示す）、シミュレーション a（図 15 に示す）を実行してシミュレートする。各処理の詳細は後述する。

【0043】

シミュレータはサーバ選択部 21 からの指示で動作してもよく、サービスモデル部 17 あるいはサーバモデル部 19 の更新時に動作してもよい。

【0044】

サーバ選択部 21 は、サーバ許容度予測処理（図 14 に示す）、限界値決定処理（図 9 に示す）を実行して、サーバ問い合わせに対し、適切なサーバを指定する。

【0045】

次に、図 3 にサービス、サーバそれぞれのシミュレーションモデルを示す。

[サーバモデルの処理フロー]

サーバモデル 30 は、2 つの待ち行列（キュー 1、キュー 2）および処理 1、処理 2 から構成される。サービスモデル 31 は処理 3 で構成される。

【0046】

最初に、サーバモデル 30 の動作から説明する。処理 2 の待ち行列（キュー 2）はサーバのネットワーク出力処理に相当し、処理 1 は CPU 処理やファイル処理など他の処理に相当する。シミュレーションではクライアントの応答はキュー 1 に入り、処理 1 を実行後、キュー 2 に入り、処理 2 の実行後、サーバの応答としてサーバモデル 30 から出て行く。キュー 1、2 は FIFO でよく、優先度付き FIFO でもよい。

【0047】

処理 1 のフローチャートを図 5 に示す。

S501でキュー1から応答*i*を取り出す。

応答*i*のサーバ処理時間 t_{p_i} は最初は未設定である。

S502で t_{p_i} が未設定であれば、S503～S506で応答*i*のタイプに応じたサーバ処理時間を t_{p_i} に設定する。 t_{p_i} はタイプ別でなくてもよい。サーバ処理総時間はサーバパラメータのサーバ処理時間とする。

S507ではサーバでの1回分の処理時間 t_{prc} を求める。 t_{prc} は単位処理時間と t_{p_i} の小さい方とする。単位処理時間は設定値でよい。

S508では応答*i*に対するサーバ処理を、シミュレーション内部での仮想時間を t_{prc} だけ進め、 t_{p_i} を t_{prc} だけ減じることで、模倣する。

S509で t_{p_i} が正であれば、S510で応答*i*をキュー1へ入れる。

S509で t_{p_i} が0または負であれば、S511で応答*i*をキュー2へ入れる。

つまり、応答*i*は t_{p_i} が0以下になるまで再びキュー1へフィードバックされ、処理1が繰り返される。

【0048】

処理2のフローチャートを図6に示す。処理2は、応答*i*に対するサーバ応答*j*についての模倣である。

S601でキュー2から応答*i*を取り出す。

S602で、(1) 応答*j*のサイズを L_j とし、(2) パケットサイズ（ネットワークに流せるパケットの最大サイズ）と L_j の小さい方を P_{ji} （クライアントとサーバ間を流れるパケットのサイズ）とし、(3) P_{ji} をサーバ送信スループット α で割ったものをサーバの送信時間 t_{ti} とし、(4) L_j から P_{ji} を差し引く。

S607で仮想時間を t_{ti} だけ進める。

S603でサービスモデルの処理3を呼び出す。処理3については後述する。

S604で L_j が正か否かを判定する。

L_j が正であれば、S605で、(1) パケットサイズ（ネットワークに流せるパケットの最大サイズ）と L_j の小さい方を P_{ji} （クライアントとサーバ間を流れるパケットのサイズ）とし、(2) P_{ji} をサーバ送信スループット α で割った

ものをサーバの送信時間 t_{ti} とし、(3) L_j から P_{ji} を差し引く。

S606で仮想時間を t_{ti} だけ進める。

【0049】

S602、S605において、パケットサイズはネットワークに流せるパケットの最大サイズMTU (maximum transfer unit) である。MTUは例えば、SNMP (simple network management protocol) を使い同じセグメントにあるルータから取得してもよく、一般に用いられている手法であるパスMTUディスカバリで求めてもよく、設定値であってもよい。

S604、S606は応答 j が複数のパケット送信からなる場合の処理である。

S606、S607でパケット送信処理を、仮想時間を進ませることで模倣する。応答 j の先頭パケットについては、サービスモデル31の処理3を呼び出す。

〔サービスモデルの処理フロー〕

次に、サービスモデル31の動作を説明する。サービスモデル31における処理3のフローチャートを図7に示す。処理3はサーバからのパケットに対するクライアントの応答を模倣する。処理3は各セッション開始時とサーバ応答後に実行される。

S701は対応するサービスの終了判定である。終了判定は例えば、設定した送受信回数に達したか否かでもよく、送受信データ総サイズが設定値に達したか否かでもよい。

S702はサービス開始であるかどうかの判定であり、開始以外ではクライアント応答であるため、S703でクライアント応答時間 (t_{crk}) だけ仮想時間を進ませる。

S704はクライアント応答 k を設定する。設定は例えば、パケットサイズや応答のタイプの設定である。

S705ではキュー1へ応答 k を入れる。

〔セッションレベルのフロー〕

図8にセッション（開始から終了まで）レベルのフローを示す。

S801でサービスモデルのパラメータからセッションの例（サービス S_i ）を

設定する。

S 8 0 2 でシミュレーション 1 が呼び出されたときの仮想時刻をセッション開始時刻 T_s として記録する。

S 8 0 3 でサービスモデルの処理 3 を呼び出す。

S 8 0 4 でサービス S_i の終了を待つ。

サービス S_i が終了したとき、S 8 0 5 で終了時刻（現仮想時刻）とセッション開始時刻 T_s の差よりサービス時間 R_i を求める。 R_i はシミュレーション結果を算出するためのサンプル値として記録する。

[サービスレベルのフロー]

図 1 2 にサービスレベルのフローを示す。

S 1 2 0 1 はシミュレーションの終了判定であり、例えば設定したシミュレーション総時間が経過したか否かでもよく、シミュレーション中、待ち行列のいずれかで定常状態の破綻が検出されたか否かでもよく、シミュレーション結果の数が設定値に達したか否かでもよい。

S 1 2 0 2 はセッションレベル処理である。

S 1 2 0 3 では仮想時間を T_i だけ進ませる。

シミュレーション 2 はサービス時間間隔 T_i 毎にセッションレベル処理を実行することで、1 セッションを模倣する。

[シミュレーションのフロー]

図 1 5 にシミュレーション a のフローを示す。

S 1 5 0 1 でサービス毎にモデルをロードする。

S 1 5 0 2 の終了判定は、図 1 2 に示す終了判定 S 1 2 0 1 と同じである。

S 1 5 0 3 で全サービスについてサービスレベル処理 a を実行する。

S 1 5 0 4 で指定サービスについての結果を記録する。

[基準値を決定するフロー]

図 9 はサーバ選択で用いる基準値を求めるフローである。

S 9 0 1 でサービス i についてセッションレベル処理を行なう。

S 9 0 2 でサービス i 結果 R_i を基準値 i とする。

サーバ選択ではシミュレーション結果が低負荷／高負荷／過負荷のいずれである

かを判定するが、そのときの結果の比較対象が基準値である。高負荷状態とはサービス処理に待ちが生じサービス時間が長くなる状態、過負荷状態とは処理できないサービスが現れる状態を意味する。基準値はサービスそれぞれを別々に1セッションだけシミュレーション評価した結果とする。

【0050】

シミュレーション結果が低負荷／高負荷／過負荷のいずれであるかの判定は例えば、結果と基準値の比率が閾値を超えたか否かでもよく、結果と基準値の差が閾値を超えたか否かでもよい。

〔サーバ許容度予測のフロー〕

図14はサーバ選択で用いるサーバ許容度予測のフローである。

S1401で $H=1$ ， $L=0$ に設定する。

S1402で指定サービスの頻度を H 倍してシミュレーションaを実行する。

S1403で低負荷状態か、高／過負荷状態かを判定する。

低負荷状態のときは、S1404で H の値を L に設定し、 H の値を2倍する。その後、S1402の処理に戻る。

一方、高／過負荷状態のときは、S1412で H の値が1か否かを判定する。

H の値が1のときは、S1405でサーバ i の許容度を0として、終了する。

H の値が1以外のときは、S1406で $(L+H)/2$ の値を M に設定する。

S1407で指定サービスの頻度を M 倍してシミュレーションaを実行する。

S1408でS1407の処理結果を判定する。

S1407の処理結果が低負荷を示しているとき、S1411で M の値を L に設定し、S1406に戻る。

S1407の処理結果が過負荷を示しているとき、S1410で M の値を H に設定し、S1406に戻る。

S1407の処理結果が高負荷を示しているとき、指定サービスの頻度の M 倍の値をサーバ i の許容度として処理を終了する。

【0051】

サーバ許容度とは例えば、サーバが高負荷あるいは過負荷状態にならない最大のサービス頻度あるいは最小のサービス間隔である。まず最初に現時点のサービ

スモデルでのサービス頻度でシミュレーションを行ない、結果が高／過負荷になるまで頻度を2倍にしていく（S1401, S1402, S1403, S1404）。

【0052】

現時点の頻度で既に高／過負荷の場合をS1412で判断し、そうであれば許容度を0とする。そうでなければ、2分探索法に基づきシミュレーション結果が高負荷となるサービス頻度を求める。

〔許容度探索の例〕

図16はサーバ許容度予測のフローにしたがって許容度探索を行う過程を示す図である。図16において横軸はサービス頻度（言い換えれば、サーバの負荷）、縦軸は所要のサービス時間である。

最初の頻度1での結果は低負荷であるので倍増していく。頻度2から頻度3としたところで、結果は過負荷となる。ここで探索処理は図14のS1403からS1412へ移り、2分探索となる。直前の頻度2と過負荷となった頻度3の中間の頻度4についてシミュレーションを行う。頻度4は低負荷であるため、さらに頻度4と頻度3の中間の頻度5でシミュレーションを行った結果、高負荷となる頻度5が求まり、頻度5をサーバの許容度とする。

【0053】

サーバ選択部21は、あるサービスについての振り分け先サーバの問い合わせを受けると、サーバそれぞれについて図14に示すフローを実行して該サービスについての許容度を求める。そして許容度と現時点での頻度との差が最も大きなサーバを振り向け先として指定する。

（実施例2）

サーバ選択について、あるサービスについての振り分け先サーバ問い合わせを受けると、サーバそれぞれについて、サーバ過負荷予測を行う。サーバ過負荷予測とはサーバへ新規サービスを振り分けた場合、過負荷状態になるかどうかをシミュレートすることである。

〔サーバ過負荷予測のフロー〕

サーバ過負荷予測のフローを図10に示す。

S1001でサーバiのモデルをロードする。

S1002でサーバiについて、後述するシミュレーションb（図11）を実行する。

S1003でシミュレーション結果Ri（サービス時間）の中で基準値iの β （所定値）倍を超えた結果Riの割合を求める。

それが γ （所定値）以下であれば過負荷にならないと予測する。（S1004）

γ より大きければ過負荷になると予測する。（S1005）

β と γ は設定値を用いてよい。

【0054】

過負荷予測でのシミュレーションbを図11により説明する。

前述した図15に示すシミュレーションaのフローとは、指定サービスについての処理が加わる点が異なる。

S1101でサービス毎にモデルをロードする。

S1102で終了か否かを判定する。

終了していない場合、S1103でシミュレーション毎に指定サービスについてシミュレーション内での開始時刻を設定する。設定は乱数を用いて決めてもよく、関数を用いてもよく、設定値としてもよい。

S1104では前述した図12に示すサービスレベル処理aに加え、後述する図13に示すサービスレベル処理bも実行する。

S1105でサービス処理bの結果を記録する。

【0055】

図13に示すサービスレベル処理bについて説明する。

S1301で指定時刻まで待機する。

S1302で指定サービスについてセッションレベル処理を行なう。

S1303でシミュレーション結果Ri（サービス時間）を出力する。

【0056】

図13のサービスレベル処理bでは指定時刻まで待機し（S1301）、1セッションだけ実行する（S1302）点が、図12のサービスレベル処理aと異

なる。

【0057】

サーバ選択部21ではサーバ過負荷予測で過負荷状態にならないサーバを選択する。

【0058】

図10に示すサーバ過負荷予測フローにおいて、S1004、S1005を省き、過負荷となった結果の割合が最も小さなサーバを選択してもよい。

(実施例3)

図17に示すようにサーバ振り分け装置を負荷分散装置3上に設け、パケット中継装置40からパケットを取り出す構成とすることもできる。

(実施例4)

各サーバの許容度をサービス振り分けの重み付けの値としたり、あるいは許容度の相対比率をサービス分配比率として、これらの値を負荷分散装置へ送出する。負荷分散装置では、サービス分配比率、あるいは重み付けラウンドロビンなどを使用して、サービスを各サーバへ振り分ける。

【0059】

前述した図14に示すサーバ許容度予測フローで各サーバの許容度 C_i を求め、 C_i を重み付けの値とする。あるいは、許容度 C_i を

$$R_i = C_i / \sum C_i$$

で比率 R_i へ変換して、比率 R_i をサービス分配比率とする。

【0060】

分配比率または重み付けの値は一度だけ求めてもよいし、定期的に図14のフローを実行して再設定してもよいし、ユーザの指示などで再設定してもよい。

(各種処理の体系)

本発明における各種の処理の体系を図18に示す。

最も下位のレベルの処理として図5、図6、図7に示すサーバモデル処理／サービスモデル処理がある。その次の段階の処理として図8に示すセッションレベル処理がある。その次の段階として図12、図13に示すサービスレベル処理がある。その次の段階として図11、図15に示すシミュレーション処理がある。

そして最上位レベルの処理として図 1 0、図 1 4 に示すサーバ過負荷予測処理／サーバ許容度予測処理がある。サーバ選択部 2 1 におけるサーバ選択処理は、この最後のサーバ過負荷予測処理／サーバ許容度予測処理の結果に基づいて実行されることになる。

【 0 0 6 1 】

【発明の効果】

本発明によれば、サーバを高負荷状態や過負荷状態にすることなく、サーバ負荷をバランスさせるサーバ振り分け装置を実現することができる。

【図面の簡単な説明】

【図 1】

本発明によるサーバ振り分け装置のネットワークシステム内の位置を示す図である。

【図 2】

本発明のサーバ振り分け装置の構成を示す図である。

【図 3】

シミュレーション・モデルを示す図である。

【図 4】

各種のパラメータの例を示す図である。

【図 5】

サーバモデルの処理 1 のフロー図である。

【図 6】

サーバモデルの処理 2 のフロー図である。

【図 7】

サービスモデルの処理 3 のフロー図である。

【図 8】

セッションレベルのシミュレーション・フロー図である。

【図 9】

1 サービスについての基準値決定フロー図である。

【図 1 0】

1サーバについてのサーバ過負荷予測フロー図である。

【図 1 1】

シミュレーション・フロー図である。

【図 1 2】

サービスレベルのシミュレーション・フロー図である。

【図 1 3】

サービスレベルのシミュレーション・フロー図である。

【図 1 4】

1サーバについてのサーバ許容度予測フロー図である。

【図 1 5】

シミュレーション・フロー図である。

【図 1 6】

許容度探索の例を示す図である。

【図 1 7】

本発明のサーバ振り分け装置の他の構成を示す図である。

【図 1 8】

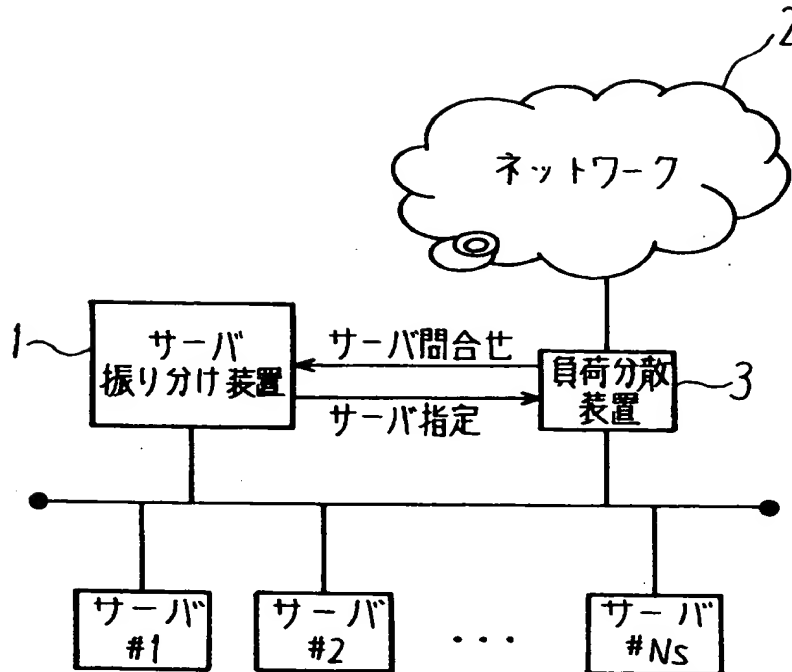
本発明における各種の処理の体系を示す図である。

【符号の説明】

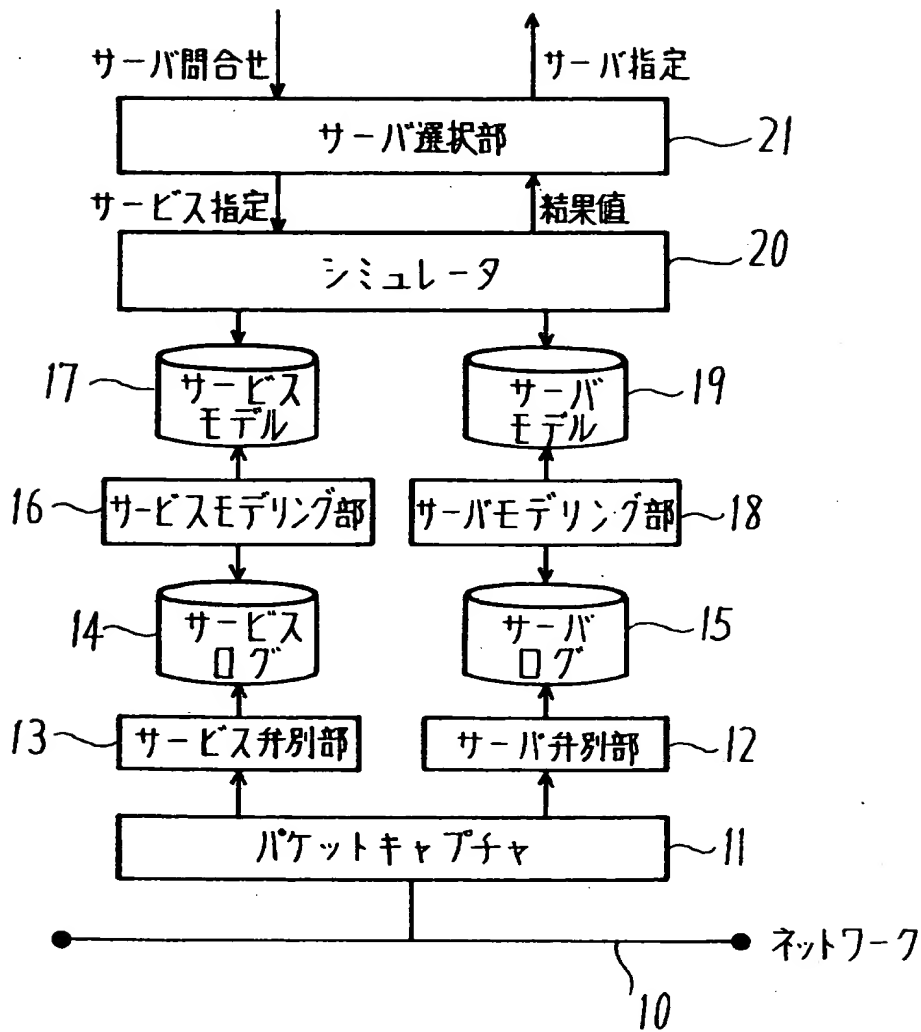
- 1 1 バケットキャプチャ
- 1 2 サーバ弁別部
- 1 3 サービス弁別部
- 1 4 サービスログ部
- 1 5 サーバログ部
- 1 6 サービスモデリング部
- 1 7 サービスモデル部
- 1 8 サーバモデリング部
- 1 9 サーバモデル部
- 2 0 シミュレータ
- 2 1 サーバ選択部

【書類名】 図面

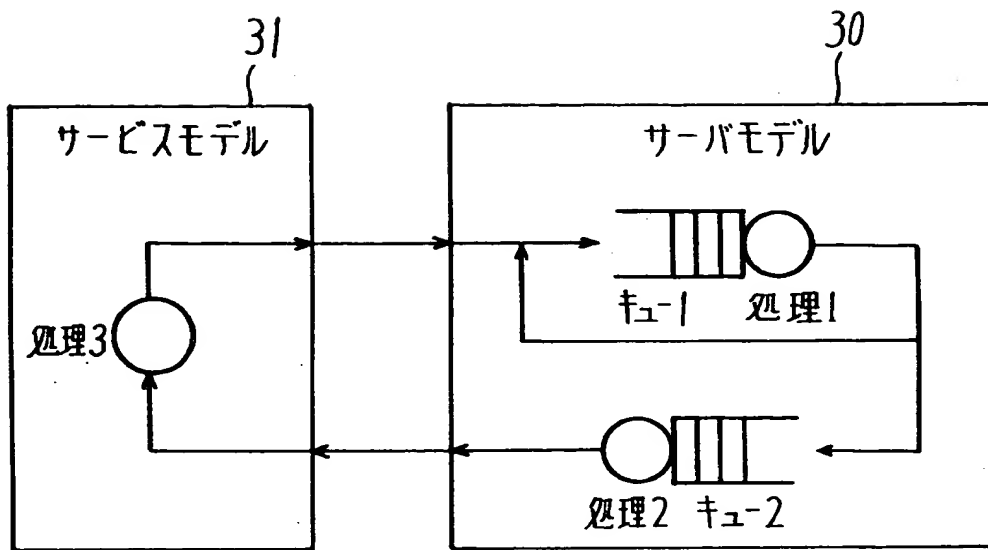
【図 1】



【図 2】

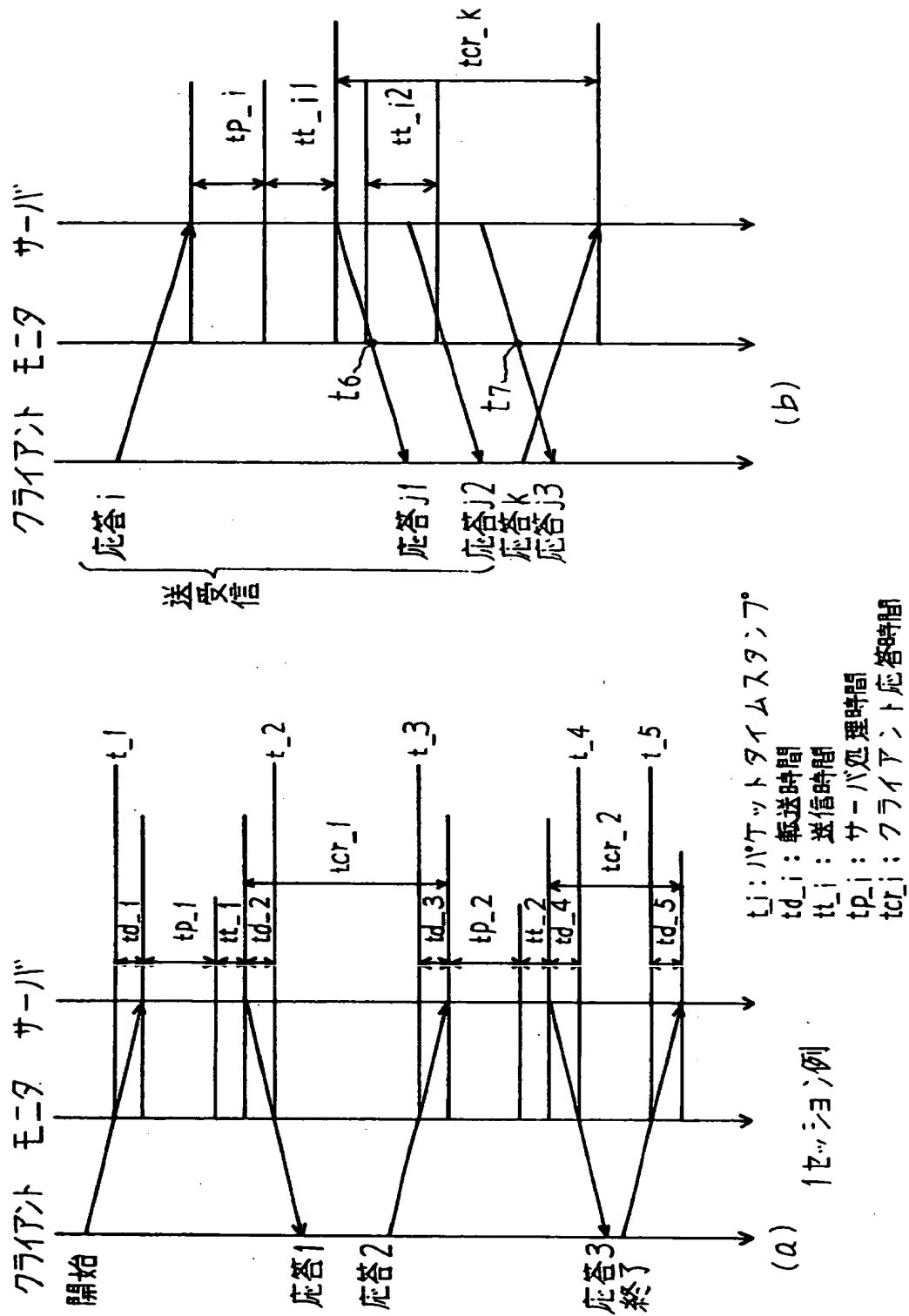


【図 3】

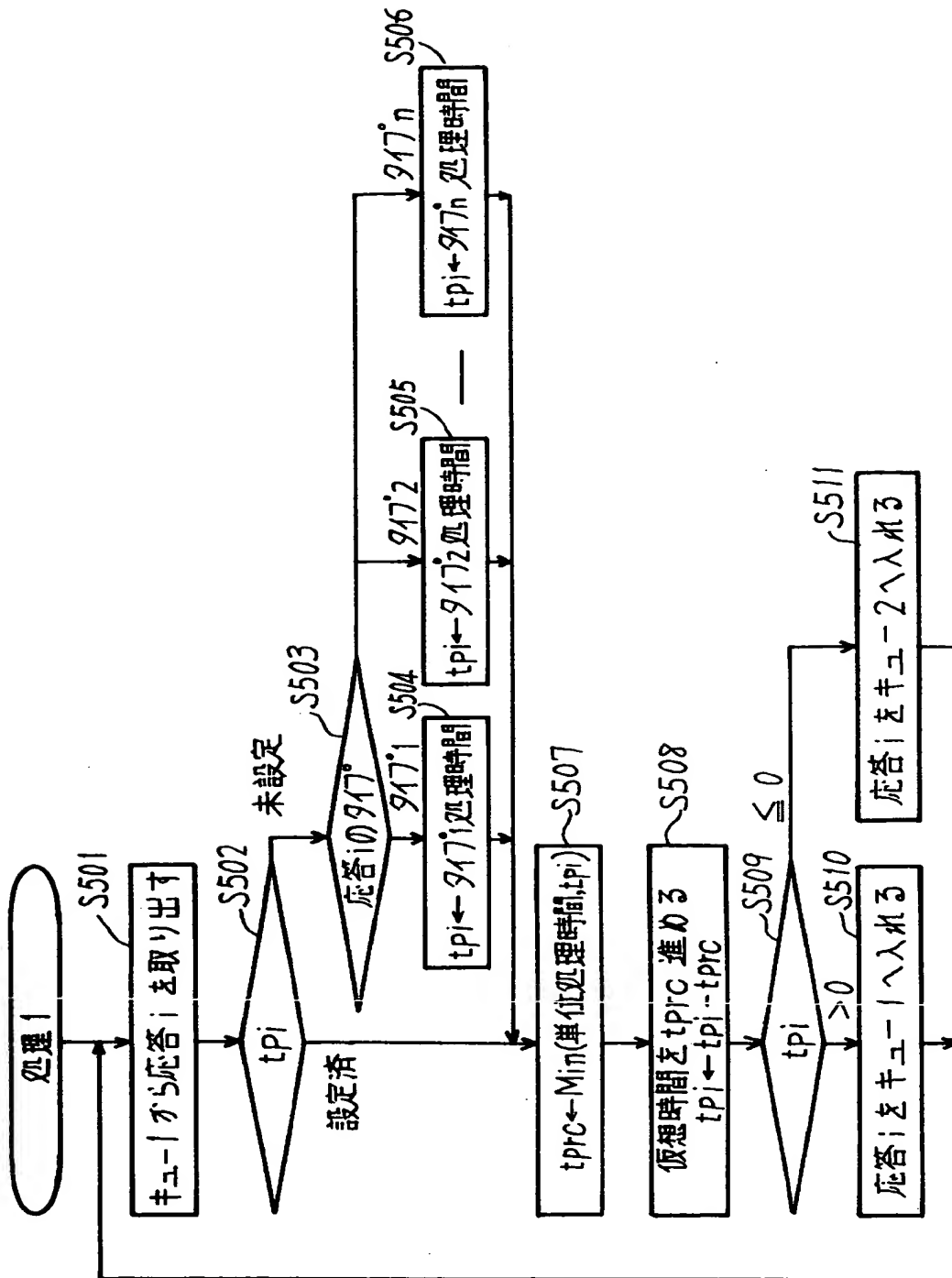


シミュレーション・モデル (1サービス、1サーバ)

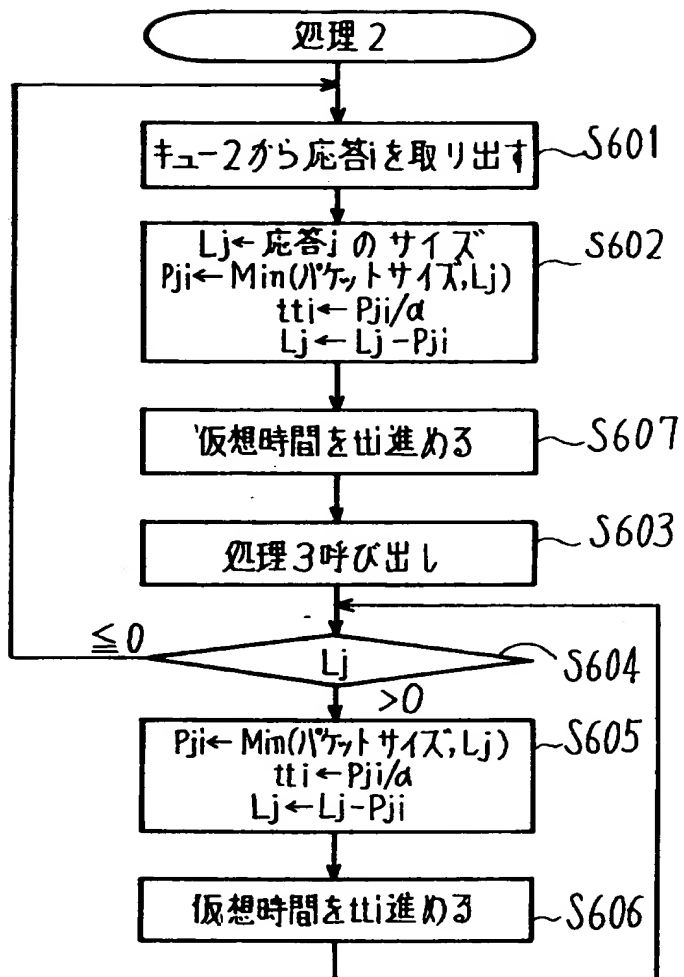
【図 4】



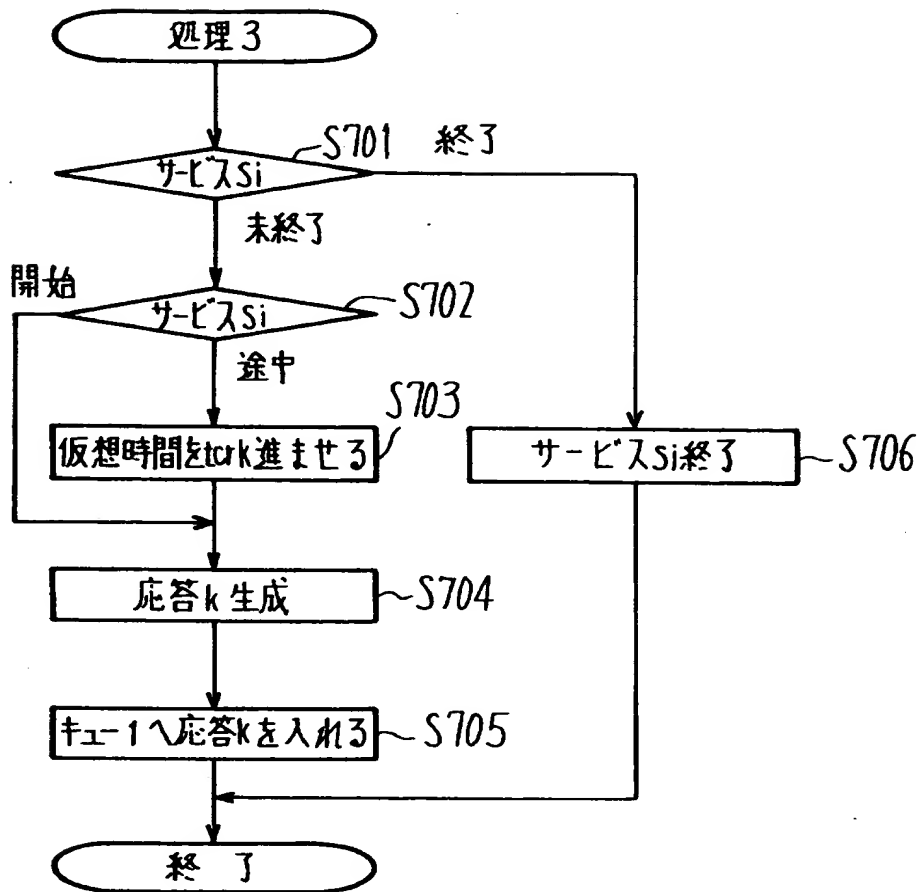
【図 5】



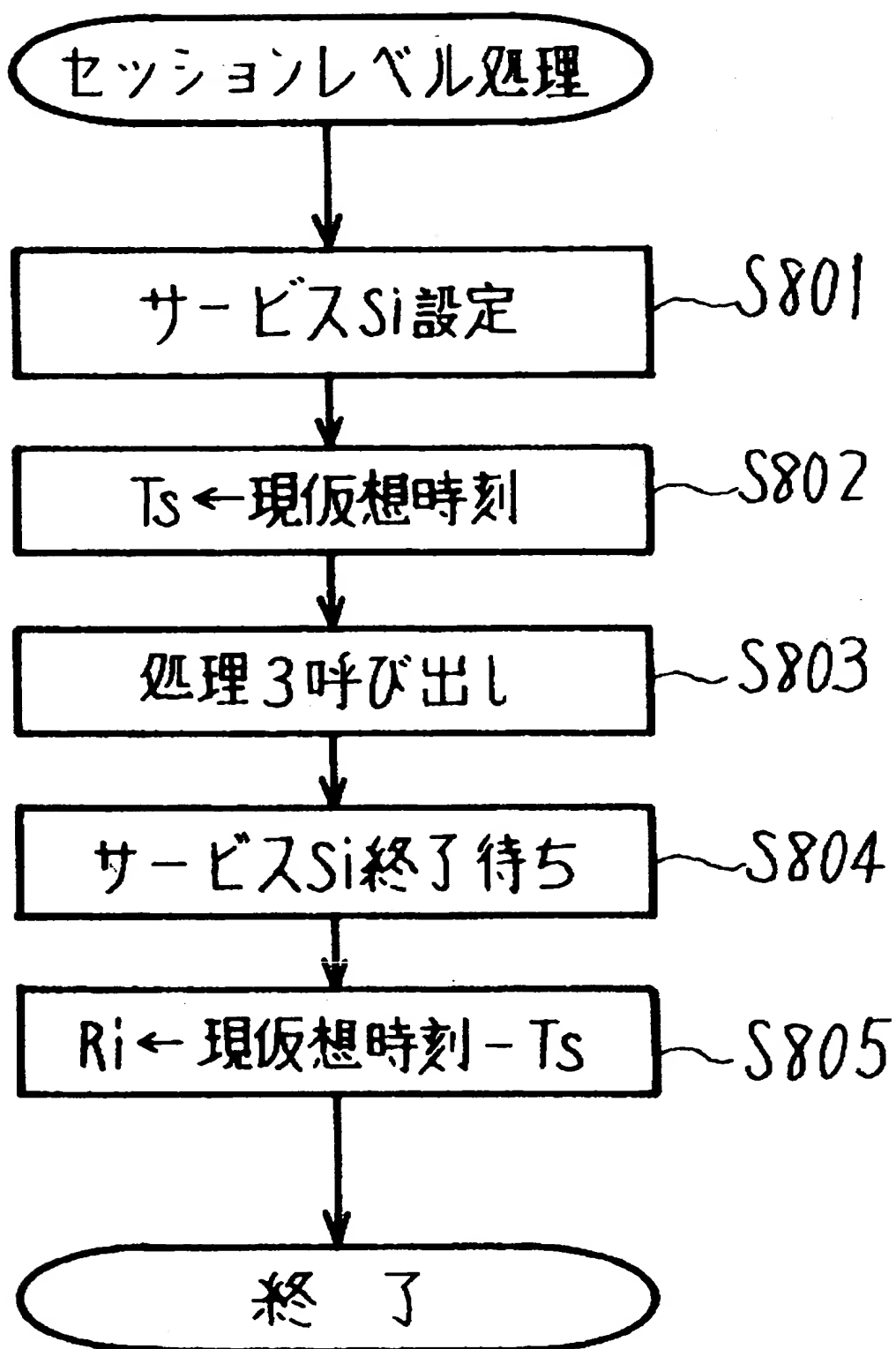
【図6】



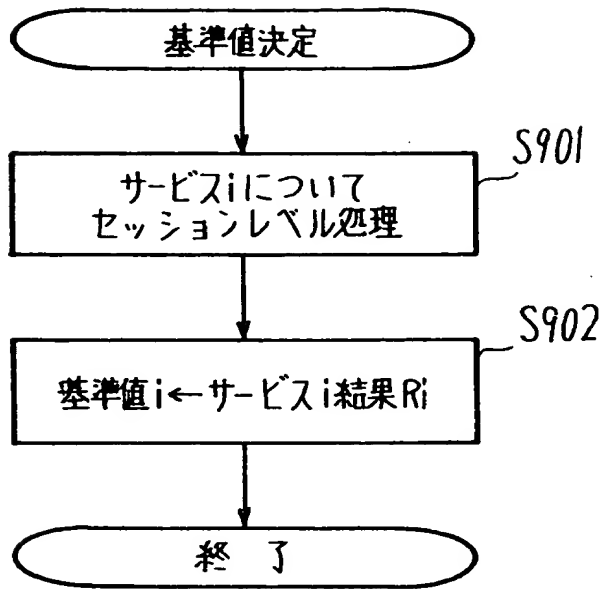
【図 7】



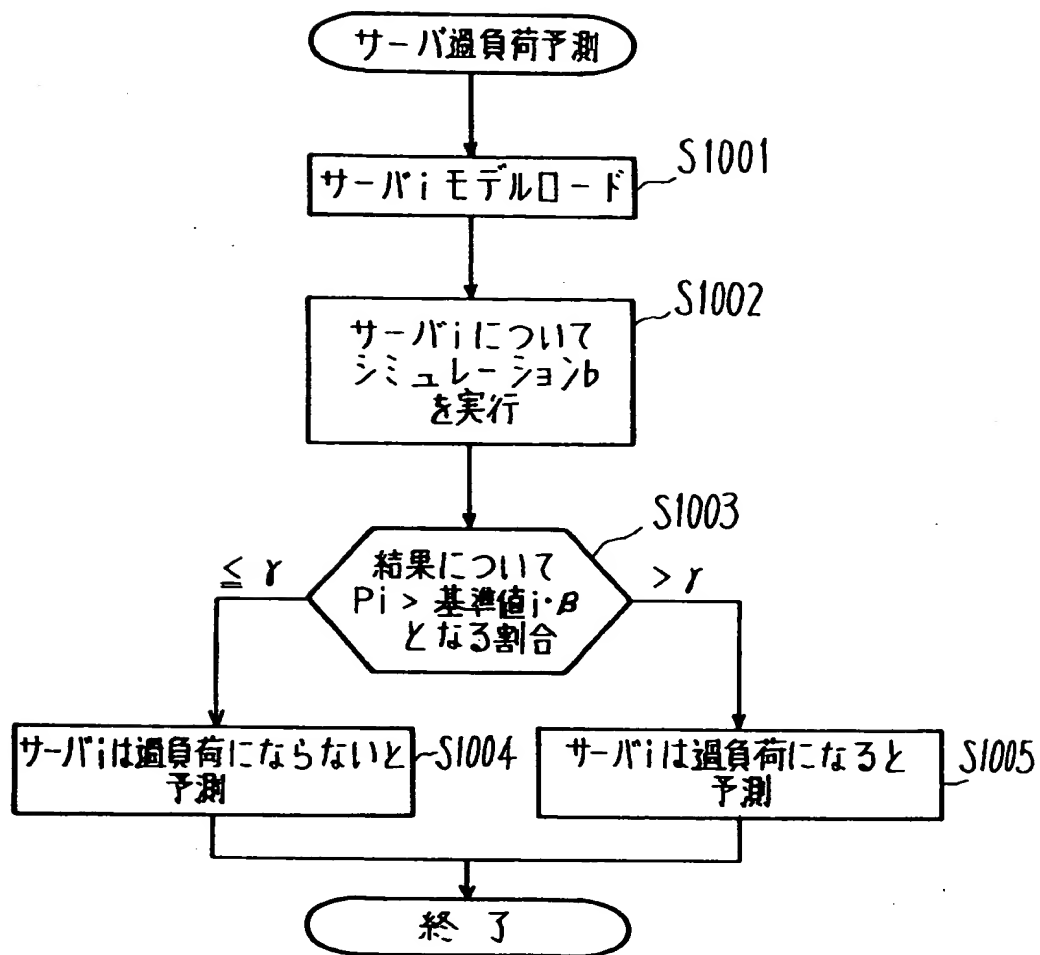
【図 8】



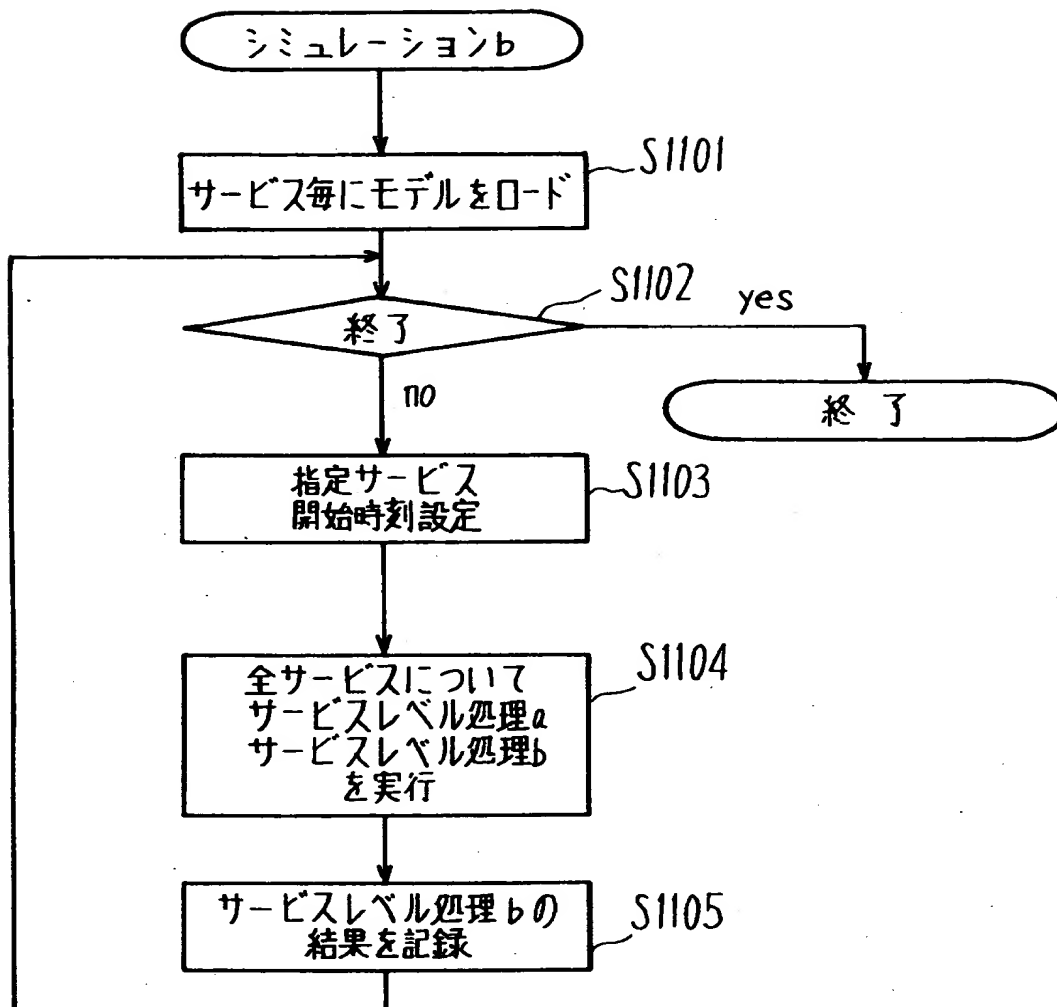
【図 9】



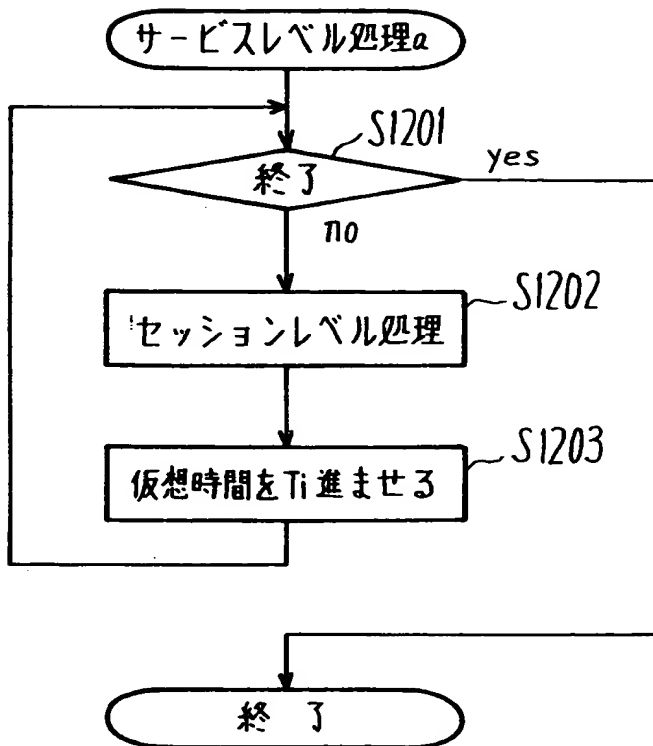
【図 10】



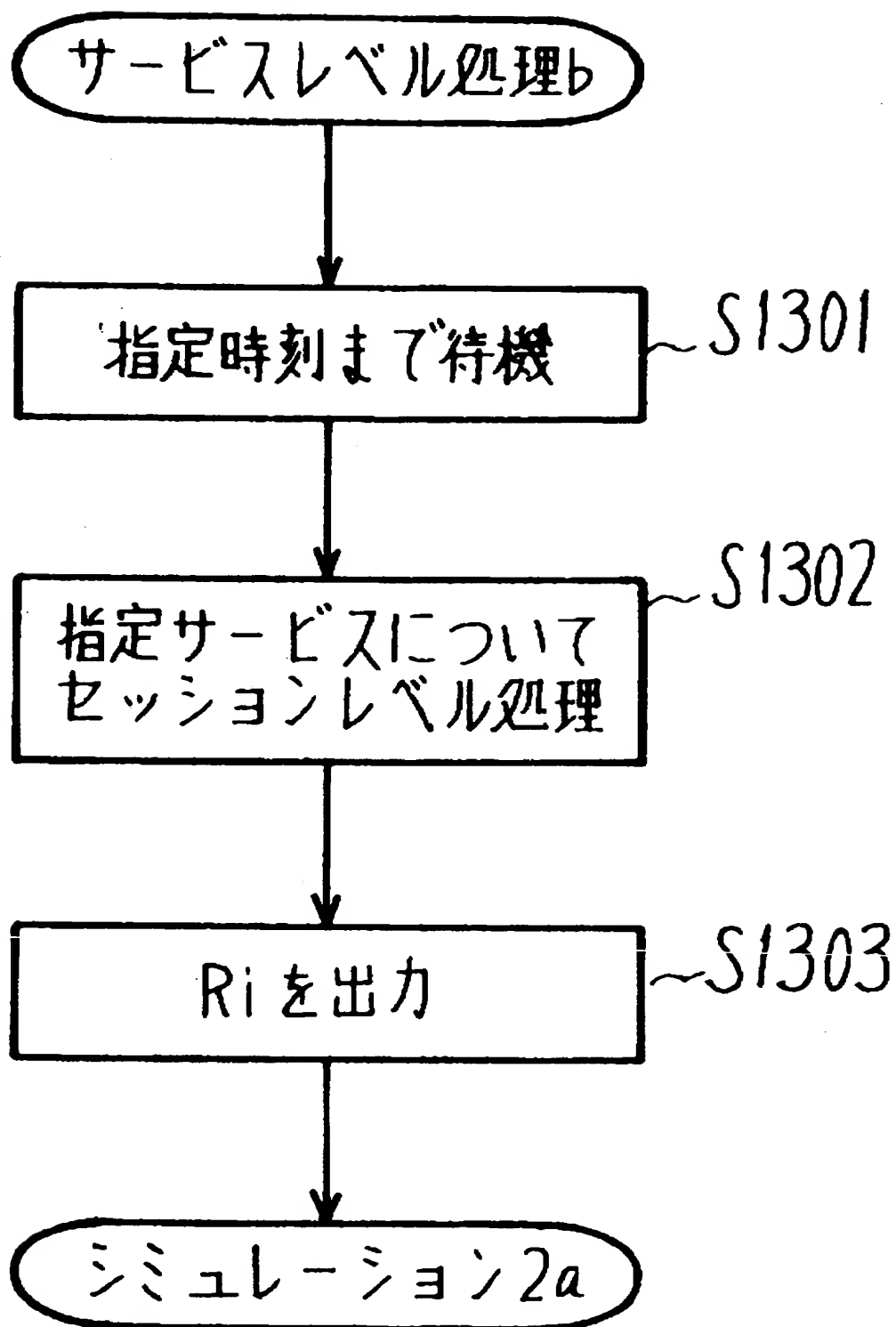
【図 1 1】



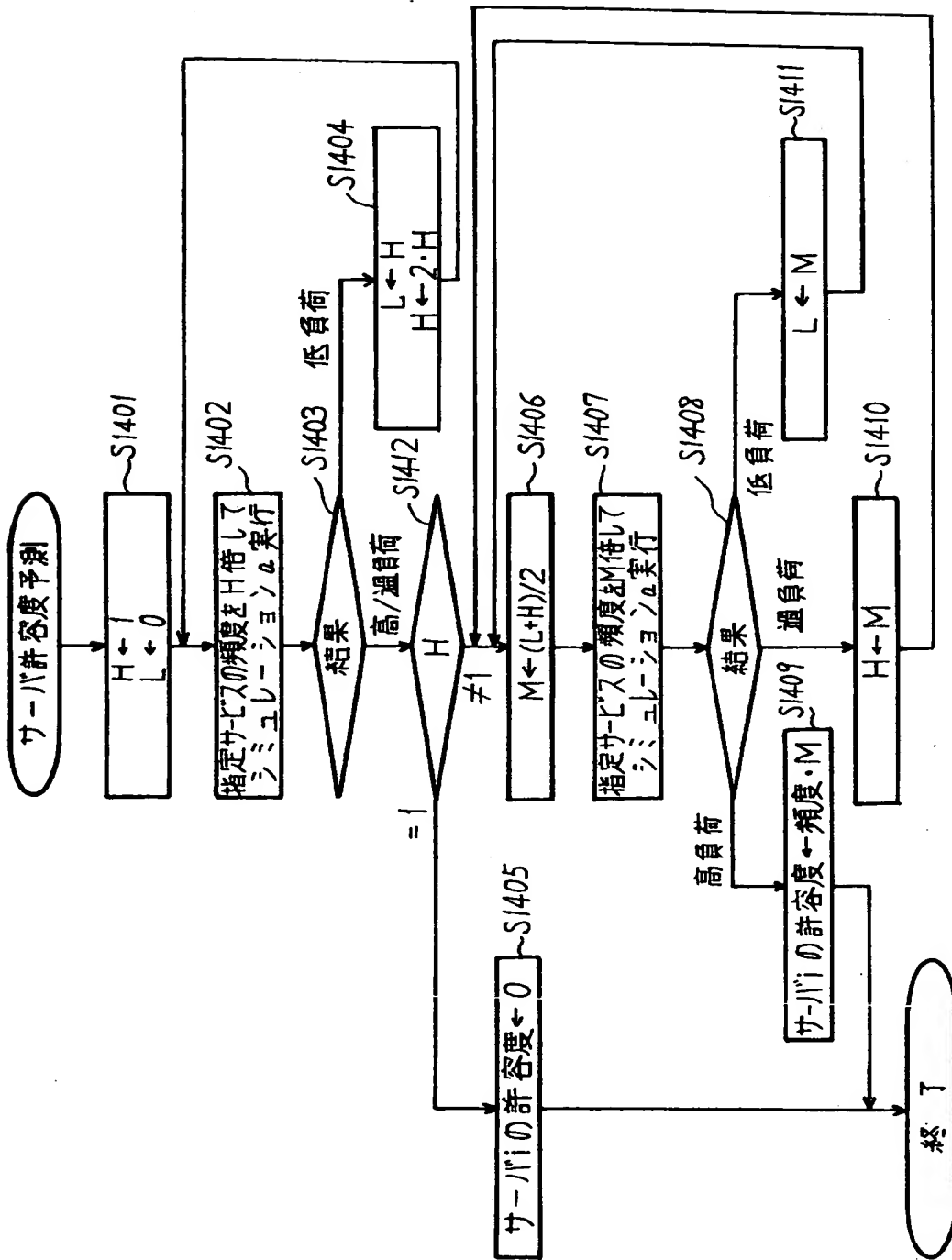
【図 1 2】



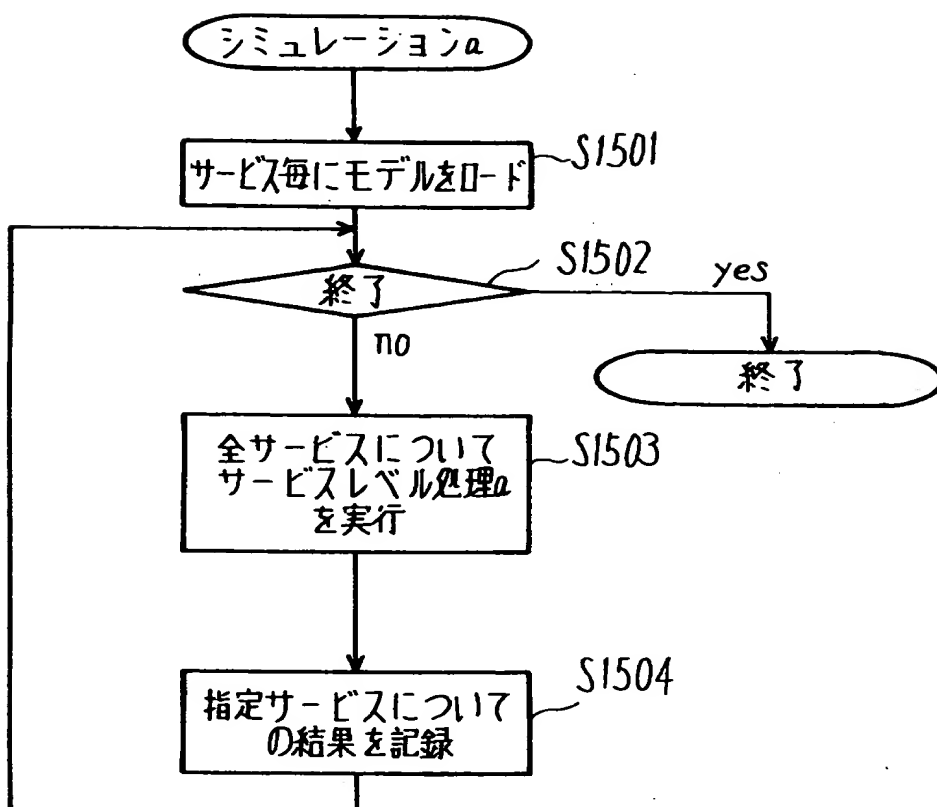
【図 13】



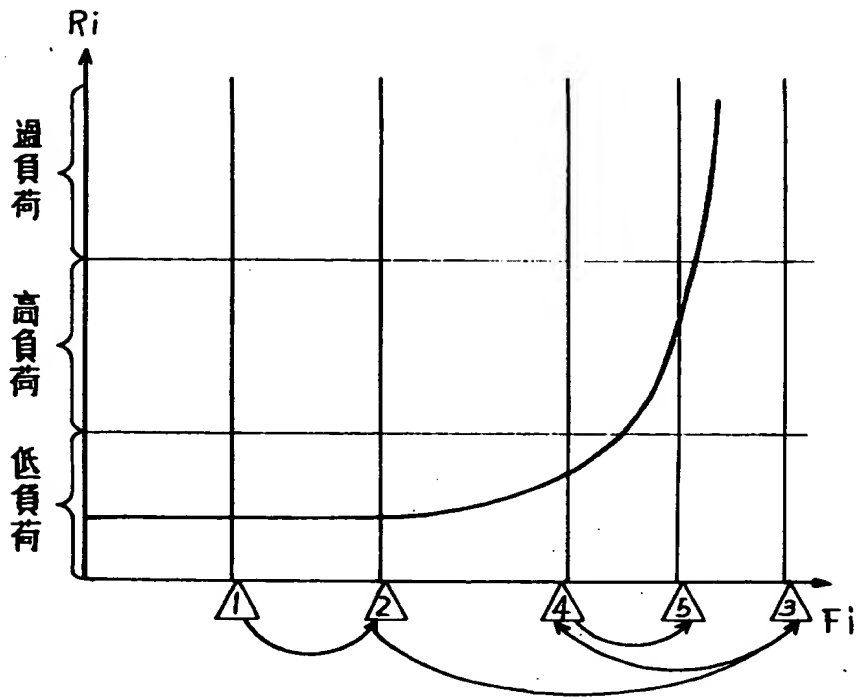
【図 14】



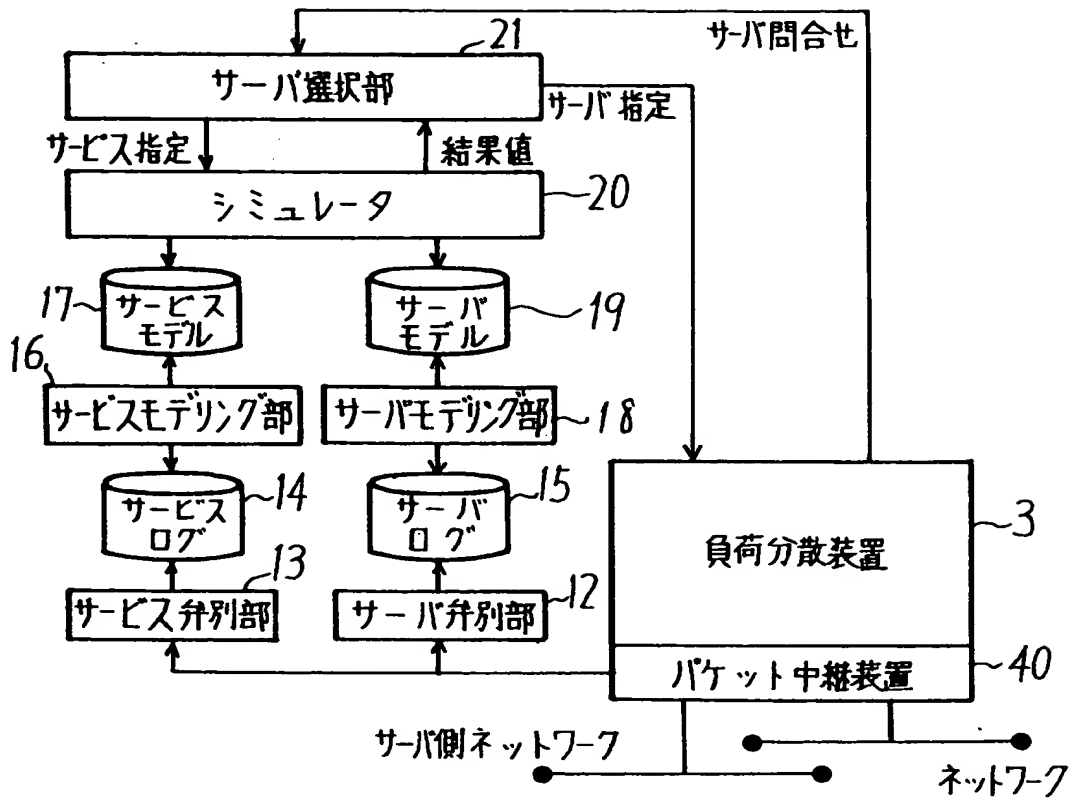
【図 15】



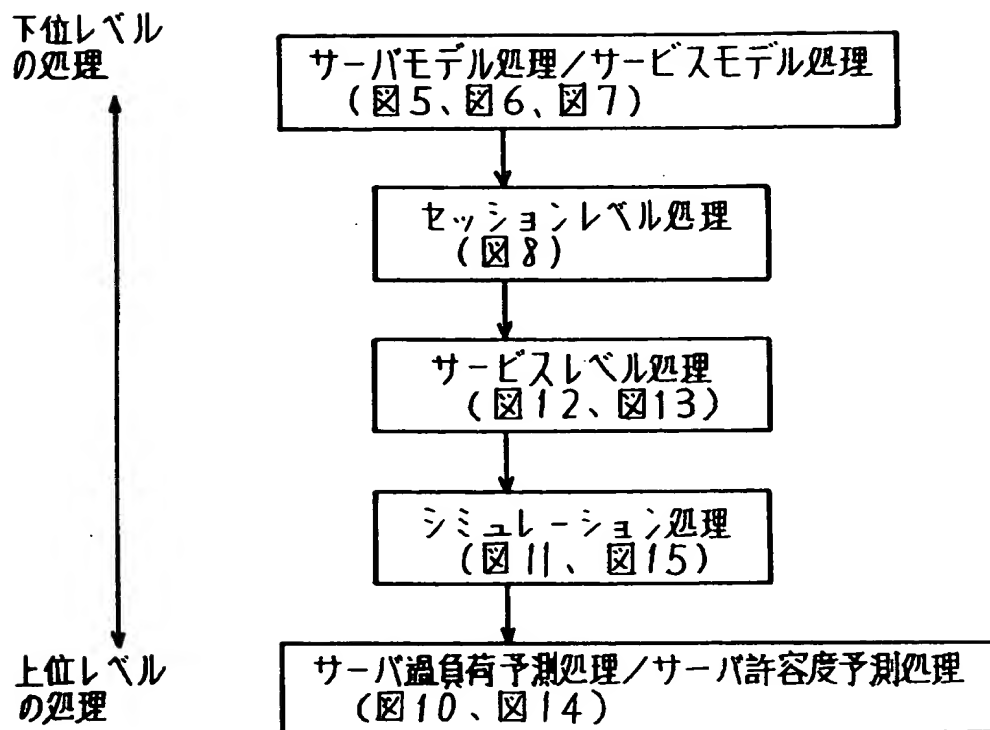
【図 1 6】



【図 1 7】



【図 18】



【書類名】 要約書

【要約】

【課題】 サーバを高負荷状態や過負荷状態にすることなく、サーバの振り分けを行う。

【解決手段】 モニタしたパケットの記録からサーバとサービスのモデルを設定し、シミュレーションを行なってサーバの負荷を予測し、予測結果からサービスの振り分け先をサーバの負荷がバランスするように決定する。

【選択図】 図 2

出 願 人 履 歴 情 報

識別番号 [000005223]

1. 変更年月日	1996年 3月26日
[変更理由]	住所変更
住 所	神奈川県川崎市中原区上小田中4丁目1番1号
氏 名	富士通株式会社